THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 273, No. 11, Issue of March 13, pp. 6508–6517, 1998
Printed in U.S.A.

Cloning, Sequencing, Expression, and Insertional Inactivation of the Gene for the Large Subunit of the Coenzyme B_{12}-dependent Isobutyryl-CoA Mutase from *Streptomyces cinnamonensis**

(Received for publication, September 16, 1997, and in revised form, December 19, 1997)

Katja Zerbe-Burkhardt‡, Ananda Ratnatilleke‡, Nathalie Philippon‡, Ashley Birch‡, Andreas Leiser‡, Jan W. Vrijbloed‡, Daniel Hess§, Peter Hunziker§, and John A. Robinson‡¶

From the Institutes of ‡Organic Chemistry and §Biochemistry, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland

Purification of the coenzyme B_{12}-dependent isobutyryl-CoA mutase (ICM) from *Streptomyces cinnamonensis* gave a protein of ~65 kDa by SDS-polyacrylamide gel electrophoresis, whose gene *icmA* was cloned using sequences derived from tryptic peptide fragments. The gene encodes a protein of 566 residues (62,487 Da), with 43–44% sequence identity to the large subunit of methylmalonyl-CoA mutase (MCM) from *S. cinnamonensis* and *Propionibacterium shermanii*. Targeted disruption of the *icmA* gene yielded an *S. cinnamonensis* mutant devoid of ICM activity. The IcmA protein is ~160 residues shorter than the large subunit of the bacterial MCMs, corresponding to a loss of the entire C-terminal coenzyme B_{12} binding domain. The sequence of the $(\beta/\alpha)_8$-barrel comprising residues A1–A400 in *P. shermanii* MCM is highly conserved in IcmA. The protein was produced in *Streptomyces lividans* and *Escherichia coli* with an N-terminal His_6 tag (His_6-IcmA), but after purification His_6-IcmA showed no ICM activity. In the presence of coenzyme B_{12}, protein from *S. lividans* and *S. cinnamonensis* of ~17 kDa by SDS-polyacrylamide gel electrophoresis could be selectively eluted with His_6-IcmA from a Ni^{2+} affinity column. After purification, this small subunit showed no ICM activity but gave active enzyme when recombined with coenzyme B_{12} and IcmA or His_6-IcmA.

Several polyketide antibiotic-producing streptomycetes have been shown to promote the interconversion of *n*- and isobutyrate. The best studied example is *Streptomyces cinnamonensis*, the producer of the commercially important polyether antibiotic monensin A (1). The interconversion of *n*- and isobutyrate occurs *in vivo* at the level of CoA[1] thioesters, as shown using a GC assay for ICM (EC 5.4.99.13) activity in cell-free extracts of *S. cinnamonensis* (2); the free acids are not substrates for the mutase. At the same time, ICM from *S. cinnamonensis* was shown to catalyze the interconversion of isobutyryl- and *n*-butyrylcarba(dethia)-CoA analogues (Fig. 1). These analogues are stable toward hydrolysis, thereby facilitating estimation of the equilibrium constant for this rearrangement, which was found to be ≈1.3 in favor of isobutyrylcarba(dethia)-CoA. The reaction catalyzed by ICM is very similar to that of the well known and widely distributed MCM (3). In both reactions, a COSCoA group migrates to an adjacent methyl, and a hydrogen atom is transferred in the reverse direction predominantly with retention of configuration (1, 4, 5).

The MCM from *S. cinnamonensis* has been cloned and sequenced (6). It was shown to be closely related in primary structure to the MCM from *Propionibacterium shermanii* (7), comprising a heterodimer with subunits of ≈65 and ≈79 kDa. The human and mouse MCMs are both homodimers with a subunit size of ≈75 kDa (8–10). Like the *P. shermanii* MCM, the *S. cinnamonensis* MCM does not catalyze the interconversion of *n*- and isobutyryl-CoA at a detectable rate (2, 6).

The structure determination of the cobalamin-binding domain of methionine synthase, a member of the methyl transferase family, revealed for the first time a protein-bound form of methylcobalamin, a vitamin B_{12} derivative (11). The cobalamin was shown bound to the protein with a histidine residue providing an axial imidazole ligand to Co^{3+}, replacing the dimethylbenzimidazole appended to the corrin ring. Stupperich *et al.* (12) had shown earlier that protein-bound cobamides can have a histidine ligand. This key histidine residue in methionine synthase is found in a motif DXHXXG, which is conserved in some (but not all) of the coenzyme B_{12}-dependent mutases (13). A similar coordination of coenzyme B_{12} by histidine was also implicated in coenzyme B_{12} bound to MCM (14).

More recently, the crystal structure of the heterodimeric MCM from *P. shermanii* was reported (15). This revealed an active site, inaccessible to solvent, that is embedded along the axis of a $(\beta/\alpha)_8$-barrel domain in the large subunit. Coenzyme B_{12} is sandwiched on one end of the β/α-barrel, between this and a C-terminal domain with a fold similar to those of flavodoxin and the cobalamin-binding domain of methylcobalamin-dependent methionine synthase (11). Apart from illuminating many important aspects of substrate and coenzyme binding to MCM, this structure also confirmed the coordination of cobalt by the histidine in the conserved DXHXXG motif within the C-terminal flavodoxin-like, coenzyme B_{12} binding domain.

We report here our efforts to purify ICM from *S. cinna-*

* This work was supported by grants from the Swiss National Science Foundation. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) U67612.

¶ To whom correspondence should be addressed: Institute of Organic Chemistry, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland. Tel.: 41-1635-4242; Fax: 41-1635-6812; E-mail: robinson@oci.unizh.ch.

[1] The abbreviations used are: CoA, coenzyme A; aa, amino acid(s); GC, gas chromatography; Hm, hygromycin ($^{R/S}$, resistant/sensitive); *hygB*, hygromycin B phosphotransferase gene, confers Hm^R; ICM, butanoyl-CoA:2-methylpropanoyl-CoA mutase; *icmA*, gene encoding the large subunit of ICM; IcmB, the small subunit of ICM; IPTG, isopropyl-β-D-thiogalactopyranoside; MCM, methylmalonyl-CoA mutase; nt, nucleotide(s); orf(s), open reading frame(s); PolIk, Klenow large fragment of *E. coli* DNA polymerase I; PAGE, polyacrylamide gel electrophoresis; Ts, thiostrepton ($^{R/S}$, resistant/sensitive); PCR, polymerase chain reaction; bp, base pair(s); kb, kilobase pair(s).

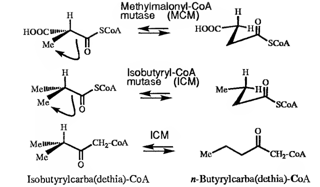

FIG. 1. **The reactions catalyzed by MCM and ICM.**

monensis, which have led to the cloning and sequencing of a gene encoding its large subunit, denoted here *icmA*. This gene was used to produce a recombinant protein in *Streptomyces lividans* and *Escherichia coli* with a His_6 tag fused to the N terminus. We also show that this recombinant protein can be used to isolate an additional small subunit of the enzyme present in *S. lividans* and *S. cinnamonensis*. This work has also allowed a comparison of the primary sequences of ICM and MCM large subunits, with unexpected implications regarding the mode of coenzyme B_{12} binding to ICM.

EXPERIMENTAL PROCEDURES

Assays

The assay used for ICM is essentially that described previously (6). Protein concentration was determined by Bradford assay (16).

Fermentation

S. cinnamonensis A3823.5 (a high yield monensin-producing strain kindly made available by Lilly (17)) was grown in 15-liter batch fermentations using a procedure described earlier (18). Cell paste (500–600 g per fermentation) could be stored at −70 °C over several weeks without substantial loss of ICM activity.

Enzyme Isolation

Buffers—Buffers were prepared as follows: buffer A, potassium phosphate (50 mM, pH 7.4) with EDTA (5 mM), dithiothreitol (1 mM), β-mercaptoethanol (0.05% v/v), and glycerol (5% v/v); buffer B, same as buffer A with phenylmethylsulfonyl fluoride (1 mM), benzamidine (1 mM), glycerol (total 20% v/v), and activated charcoal (20 g/liter); buffer C, same as buffer A with KCl (1.0 M); buffer D, same as buffer A with KCl (0.1 M) and Tris-HCl (0.1 M); buffer E, same as buffer A but with 20% glycerol; buffer F, Tris-HCl (250 mM, pH 8.3), glycine (1.92 M); buffer G, Tris-HCl (100 mM, pH 8.2), NaCl (1.0 M), $CaCl_2$ (2.0 mM), and MeCN (10%); buffer H, same as buffer A with Tris-HCl (0.15 M); buffer I, sodium acetate (0.1 M, pH 4) and NaCl (0.5 M); buffer J, Tris-HCl (100 mM, pH 8) and NaCl (0.5 M); buffer K, potassium phosphate (50 mM, pH 7.4), KCl (300 mM), glycerol (5% v/v), imidazole (20 mM), β-mercaptoethanol (0.05% v/v), benzamidine (1 mM), and phenylmethylsulfonyl fluoride (1 mM); buffer L, same as buffer K except imidazole (300 mM); buffer M, potassium phosphate (50 mM, pH 7.4), β-mercaptoethanol (0.05% v/v), dithiothreitol (1 mM); buffer N, potassium phosphate (50 mM, pH 7.4), KCl (150 mM).

Affinity Chromatography—A vitamin B_{12} affinity column (19, 20) was prepared as follows. Vitamin B_{12} (130 mg) in aqueous HCl (0.5 M, 46 ml) was stirred at 37 °C for 3 h. The solution was neutralized with aqueous NH_3 and applied to a column of Alumina N (2 × 36 cm, ICN, Germany). After eluting unchanged vitamin B_{12}, partially hydrolyzed cobalamins were eluted with aqueous NH_3 (0.2 M). After lyophilization, these were applied in water to Q-Sepharose (1.6 × 20 cm, Pharmacia Biotech Inc.), and monocarboxylic acids were separated from di- and tricarboxylic acids by elution with a gradient from 0.2 M triethylamine, pH 11, to 0.2 M triethylamine, 0.5 M acetic acid, pH 1. TLC on cellulose plates (eluting with *sec*-butyl alcohol/acetic acid/water (127:1:50)) was used to monitor this separation (R_f (B_{12}) = 0.5, R_f (monocarboxylic acids) = 0.6). Fast atom bombardment-mass spectrometry of the monocarboxylic acid fraction gave m/z = 1356.3 (M^+). The monocarboxylic acids (16 mg) were then coupled over 16 h to EAH-Sepharose (5 ml, Pharmacia) using *N*-ethyl,*N'*-(3-dimethylaminoisopropyl)carbodiimide and protocols recommended by the manufacturer (Pharmacia). The gel was then washed with water, with buffer I, then buffer J, and finally with water.

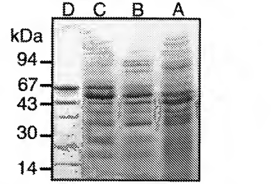

FIG. 2. **Coomassie Blue-stained SDS-PAGE gel of protein obtained after the purification of ICM from *S. cinnamonensis* (see Table I): *lane A*, after DEAE; *lane B*, after Superdex; *lane C*, after preparative gel electrophoresis; *lane D*, after vitamin B_{12} affinity chromatography.** Positions of molecular mass standards are shown.

Cell Disruption and Ammonium Sulfate Fractionation—Ultrasonic disruption of cell paste (≈500 g) in buffer B (750 ml) was carried out over 15 min at 4–10 °C, and then solids were removed by centrifugation (27,500 × *g* for 45 min). Ammonium sulfate was added to the supernatant to 35% saturation at 4 °C and pH 7.5. After centrifugation (47,800 × *g*, 1 h) and filtration through glass wool, ammonium sulfate was added to 75% saturation. Centrifugation (47,800 × *g* for 1 h) afforded a protein pellet (Table I), which could be stored at −80 °C over several months.

Chromatography on DEAE-Sepharose—Protein from above (≈5 g wet weight) was applied to DEAE-Sepharose (2.6 × 45 cm, Pharmacia) equilibrated with buffer A and eluted with a gradient (0–100% buffer C) over 500 ml at a flow rate of 3 ml/min. ICM appeared in the eluate at ≈92–100% buffer C. The active fractions containing ≈300 mg of protein were concentrated by ultrafiltration (Centriprep-10, Amicon).

Q-Sepharose—Protein from the foregoing step (in 3 batches, total ≈1.0 g) was applied to a column of Q-Sepharose (2.6 × 15 cm, Pharmacia) pre-equilibrated with buffer A and eluted with a gradient (0–50% buffer C) over 540 ml at a flow rate of 4 ml/min. ICM eluted at ≈34–42% buffer C, whereas MCM eluted at ≈47–50% buffer C. The fractions containing ICM were concentrated by ultrafiltration (Centricon-10, Amicon).

Gel Filtration—The active protein from the preceding step (6 × 20-mg batches) was applied to a Hiload 16/60 Superdex-200 column (≈120-ml bed volume, Pharmacia) pre-equilibrated with buffer D and eluted (0.3 ml/min) with buffer D. ICM eluted at ≈69–75 ml.

Preparative Gel Electrophoresis—Continuous preparative native gel electrophoresis was performed with a model 491 Prep-Cell (Bio-Rad). Protein from the preceding step (20 mg) was applied in buffer E (3 ml) to a gel comprising a stacking layer (5.25% acrylamide) followed by the fractionating gel (8% acrylamide). Electrophoresis (at 40 mA) was performed using buffer F as running buffer at 4 °C. Proteins eluted from the gel were diverted to a fraction collector.

Affinity Chromatography—Protein from the previous step (2.5 mg) was applied to affinity resin (1 ml, see above) in buffer A. The column was then eluted with a KCl gradient (10–100% buffer C). ICM appeared from the column at ≈0.2 M KCl. This fraction showed a major protein band on SDS-PAGE (Fig. 2), with apparent mass of ≈65 kDa and several minor components of lower mass.

Peptide Sequencing

The ICM-containing protein from above (≈200 μg) was electrophoresed by SDS-PAGE (12%, 10 × 10-cm gel), electroblotted onto a cationic polyvinylidene difluoride membrane (Immobilon CD, Millipore), and visualized by negative staining (Quickstain, Zoion Research). The membrane spot containing the adsorbed protein (of ≈65 kDa) was cut out and incubated in buffer G (10 μl) with trypsin (0.3 μg, Promega) for 15 h at 37 °C. Free peptides were washed from the membrane with 10% aqueous trifluoroacetic acid (1 μl) and 10% aqueous MeCN with 0.1% trifluoroacetic acid (10 μl). The peptides were analyzed by microbore high pressure liquid chromatography (C_8 column, 300 ≈ pore size, 1 × 250 mm, Vydac) eluting with a gradient of 2–80% v/v MeCN in water,

with 0.05% trifluoroacetic acid. The eluate was split; 90% was collected, and the remainder was analyzed by electrospray-mass spectrometry. Selected fractions were subjected to sequence analysis using the automated Edman method (Applied Biosystems 477A sequencer) (see legend to Fig. 4).

PCR Amplification of an icmA Gene Fragment

The following oligonucleotides were designed for the PCR using peptide sequences PAYKPLSV and TQTAGVSL determined as described above (see Fig. 4): KB1, 5′-CCGGC(G/C)TACAAGCCCCTCT-CGG-3′; KB2, 5′-CAGCGA(G/C)ACGCCCGCTGTCTGCGT-3′. PCR amplification was carried out in the recommended buffer (Vent™; 100 μl) under mineral oil containing the following: dNTPs (200 μM), KB1 and KB2 primers (0.5 μM), *S. cinnamonensis* DNA (10 ng, see below), Vent™ DNA polymerase (2 units). The reactions were performed using a Perkin-Elmer 480 thermal cycler as follows: 1 min at 94 °C, 2 min at 55 °C, and 2 min at 72 °C. After 30 cycles, the PCR product (≈310 bp) was gel-purified and cloned into *Sma*I-digested M13mp18 (21) and sequenced. The *S. cinnamonensis* DNA was prepared by digesting genomic DNA (10 μg) with *Eco*RI, *Bam*HI, *Pst*I, *Bgl*II, and *Sma*I, precipitating with EtOH, and redissolving in TE (100 μl). DNA (1 μg) was then denatured in water (32 μl) by addition of 4 M NaOH, 4 mM EDTA (8 μl) for 10 min (22). After addition of sodium acetate (3 M, 7 μl, pH 4.8) and water (4 μl), the DNA was precipitated with EtOH and redissolved in water (100 μl).

Gene Cloning and Sequencing

General DNA manipulation was performed in *E. coli* (23, 24) and in *Streptomyces* according to Ref. 25. *S. cinnamonensis* genomic DNA was partially digested with *Sau*3A, fractionated by sucrose density centrifugation, and fragments of ≈15 kb were ligated between the *Bam*HI/*Sal*I sites in λEMBL4 DNA (26). Recombinant λEMBL4 clones were isolated from this library by plaque hybridization using the ≈310-bp PCR product as probe DNA. A 7.6-kb *Bam*HI fragment containing the *icmA* gene was isolated from one λ clone and ligated into *Bam*HI-digested pUC18, to afford pOCI602 (Fig. 3). A 3.75-kb *Bam*HI/*Bgl*II fragment from this region was ligated in both orientations into *Bam*HI-cut pUC18 to afford pOCI609 and pOCI613. Similarly, a 4.1-kb *Bam*HI/*Sph*I fragment from pOCI602, after end-filling with PolIk, was ligated into *Sma*I-digested pUC18 to afford pOCI611 and pOCI612 (Fig. 3). The plasmids pOCI609, pOCI611, pOCI612, and pOCI613 were used to determine the 4.3-kb nt sequence shown in Fig. 4, on both DNA strands, by the dideoxy method (27) using dye terminator chemistry (Perkin-Elmer). Sequence information has been submitted to the EMBL-GenBank™ data base, accession number U67612.

Insertional Inactivation of IcmA

A *Bst*EII/*Bam*HI fragment containing the *icmA* gene was isolated from pOCI611, filled-in with PolIk, and cloned into the *Sma*I site of a pUC18 derivative lacking an *Sac*I site, to give pOCI641. The pUC18 derivative was made by digesting pUC18 with *Sac*I, digesting with PolI, and re-ligating. In this way, a unique *Sac*I site is available close to the center of *icmA* (*cf.* Fig. 3 and Fig. 6). A *Bgl*II/*Bam*HI fragment including the *hygB* gene was isolated from pIJ963 (28) and, after filling ends with PolIk, was inserted into the PolI-digested *Sac*I site in pOCI641 to give pOCI642. In this way, *hygB* is inserted into the unique *Sac*I site in *icmA*. The disrupted *icmA* gene was then recovered by partial digestion with *Eco*RI/*Hin*dIII, end-filled with PolIk, and cloned into the PolIk-filled *Bam*HI site in pGM160 (29), to give pOCI643. However, the resulting plasmid could not be introduced into *S. cinnamonensis*. To improve stability of the construct in *Streptomyces*, the entire *E. coli* sequences were deleted by digesting pOCI643 with *Eco*RI/*Hin*dIII, end-filling with PolIk, and religation. The resulting plasmid was passaged through *S. lividans* 1326, then denatured using the procedure of Oh and Chater (30), and used to transform *S. cinnamonensis* to Hm^R and Ts^R by selection on R5 agar plates (25) at 30 °C. The transformant was then grown in liquid YEME medium (25) with Hm at 39 °C and then plated onto R5 agar medium with Hm at 39 °C. After sporulation, relicaplating separately to R5 with Ts, and R5 with Hm, gave several Ts^S Hm^R transformants, one of which was selected for further investigation. Southern blotting (Fig. 6) demonstrated that the *icmA* gene in this transformant had been inactivated by insertion of the *hygB* gene. No ICM activity was observed in cell extracts of the mutant grown in the usual way.

Production of IcmA and His_6-IcmA

The *icmA* gene was amplified by PCR using the following primers (*Nde*I and *Bam*HI sites are underlined): KB3, 5′-CCAT<u>GGATCC</u>TCA-GAACCCGGGCGGCTCCAC-3′; KB4, 5′-GCTA<u>CATATG</u>GACGCTGAC-GCGATCGAGGAAGG-3′.

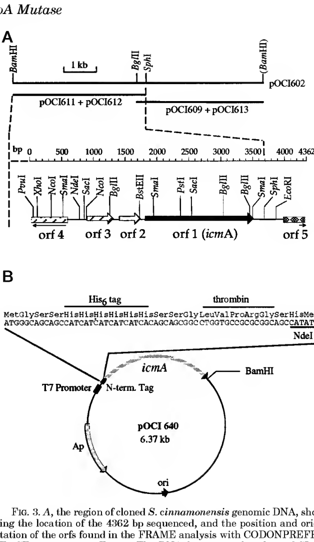

FIG. 3. *A*, the region of cloned *S. cinnamonensis* genomic DNA, showing the location of the 4362 bp sequenced, and the position and orientation of the orfs found in the FRAME analysis with CODONPREFERENCE (see text and Fig. 5). The DNA fragments cloned in pOCI602/609/613/611/612 are indicated. The *Bam*HI site in *brackets* originates from the multiple cloning site in λEMBL4. *B*, the expression plasmid pOCI640 used to produce His_6-IcmA in *E. coli* (see text). This construct was made by cloning *icmA* between the *Nde*I and *Bam*HI sites in pET14b (Novagen). Shown is the N-terminal tag sequence including His_6, the thrombin cleavage site, and the first residue (Met) of *icmA*.

A PCR was performed with primers KB3 and KB4 (1.0 μM), dNTPs (440 μM), and denatured template DNA (pOCI612, 0.5 μg), using the conditions described above. The PCR product was gel-purified, digested with *Nde*I and *Bam*HI, and cloned between the *Nde*I/*Bam*HI sites in the plasmid pET3a (31) to give plasmid pOCI614. The correctness of the *icmA* nt sequence in this plasmid was proven by DNA sequencing. The gene was then excised by digestion with *Nde*I and *Bam*HI and cloned between the *Nde*I/*Bam*HI sites in the plasmids pIJ4123 (32) and pET14b (Novagen) to afford plasmids pOCI633 and pOCI640, respectively.

Plasmid pOCI614 was introduced into *E. coli* BL21(DE3)pLysS (31). After growth in LB medium (200 ml) with ampicillin and chloramphenicol at 30 °C and induction at A_{600} = 0.7 with 0.4 mM IPTG, cells were shaken for a further 4 h at 30 °C, then collected, washed, and sonicated in buffer A. After removal of cell debris by centrifugation, IcmA was purified by chromatography on Q-Sepharose (Pharmacia) with buffer A and a gradient of 0–100% KCl (0.5 M). After dialysis against buffer A, gel filtration (TSK G3000SW) with buffer A afforded IcmA (4.3 mg) which was homogeneous by SDS-PAGE and gave the expected N-terminal amino acid sequence.

Plasmid pOCI640 was introduced into *E. coli* BL21(DE3)pLysS. After growth in LB medium (4 liters) with ampicillin and chloramphenicol at 30 °C and induction at A_{600} = 0.6 with IPTG (0.1 mM), cells were shaken for a further 4 h at 30 °C, then collected, washed, and sonicated in buffer K. After removal of cell debris by centrifugation, the protein

```
   <-- L  R  Q   I  A  E   S  L  L   Q  Y  R  D   A  H  Y   E  A  P   T  C  G  D   V  I  H   P  I  A   D  P  E  L
   1 TGAGGCGCTG GATCGCCTCG GAGAGCAGCT GGTAACGGTC CGCGTGGTAC TCGGCCGGGG TGCAGCCGTC CACGATGTGC GGGATCGCGT CGGGCTCGAG
       I  V  L   A  P  A   D  G  I  A   D  A  F   T  D  V   W  S  R  Y   A  E  A   S  A  A   G  G  A  S   H  Q  G
 101 GATCACCAGG GCGGGGGCGT CGCCGATCGC GTCGGCGAAC GTGTCCACCC AGCTCCGGTA GGCCTCCGCA CTGGCCGCGC CGCCCGCGGA GTGCTGACCG
       C  D  R   H  P  I  N   Y  A  V   L  V  A   T  R  D  E   K  V  A   G  R  T   A  K  A  V   D  P  A   P  D  D
 201 CAGTCGCGGT GCGGGATGTT GTACGCGACG AGTACGGCGG TGCGGTCCTC CTTGACCGCG CCCCGCGTCG CCTTCGCGAC GTCGGGCGCC GGATCGTCCC
      G  A  P  W   V  A  M   A  R  E   S  I  R  R   L  T  E   A  D  E   A  R  G  Q   E  E  W   Q  R  V   Q  R  A  A
 301 CGGCCGGCCA CACGGCCATG GCCCGTTCGG AGATGCGCCT GAGCGTCTCG GCGTCCTCGG CGCGGCCCTG TTCCTCCCAC TGCCTGACCT GGCGCGCGGC
       P  S  D   P  D  V   W  F  T  G   A  P  P   G  A  S   A  T  A  P   K  R  V   A  A  E   E  K  T  G   D  S  G
 401 GGGGCTGTCG GGGTCGACCC AGAAGGTGCC GGCGGGGGGC CCGGCGCTCG CGGTGGCGGG CTTGCGCACG GCCGCCTCCT CCTTCGTGCC GTCGGACCCC
       P  D  S   S  S  C  G   A  P  L   G  L  A   A  L  A  A   L  A  T   F  T  R   L  L  R  H   M     <- orf 4
 501 GGGTCTGAGG AGGAGCAGCC TGCCGGGAGC CCGAGGGCGG CGAGGGCCGC GAGTGCCGTG AACGTGCGGA GCAGCCGGTG CATCCAGCCC CCTTGGGCGA
 601 TGGTGACAGT GACGGTCAGT CAGCCCGGCA ATCGTTACAT AAAGGACTAT TCAAGCTCTT GTGCCACACC GCCTCCGGTG CCGAGCGCGA ACCCGGCGGG
 701 CACCAGAGCC CCGCCGCGGC CGCGGAGCCG TACGTACGAC CGAATTGCGA GACGGGGCTG ACCACCATAT GACCGGCGGG TAAGGTCGAT GCCGTGCCGA
                                                                                          orf 3 ->       V  P  K
 801 AGCCGCTCAG CCTCCCCTTC GATCCCATCG CCCGCGCCGA CGAGCTCTGG AAGCAGCGCT GGGGATCGGT CCCGGCCATG GGCGCGATCA CCTCGATCAT
       P  L  S   L  P  F   D  P  I  A   R  A  D   E  L  W   K  Q  R  W   G  S  V   P  A  M   G  A  I  T   S  I  M
 901 GCGGGCGCAC CAGATCCTGC TCGCCGAGGT CGACGCGGTC GTCAAGCCGT ACGGACTGAC CTTCGCGCGC TACGAGGCGC TGGTGCTCCT CACCTTCTCG
       R  A  H   Q  I  L  L   A  E  V   D  A  V   V  K  P  Y   G  L  T   F  A  R   Y  E  A  L   V  L  L   T  F  S
1001 CAGGCCGGCG AGTTGCCGAT GTCGAAGATC GGCGAGCGGC TCATGGTGCA CCCGACCTCG GTCACGAACA CCGTGGACCG CCTGGTGAAG TCCGGCCTGG
      Q  A  G  E   L  P  M   S  K  I   G  E  R  L   M  V  H   P  T  S   V  T  N  T   V  D  R   L  V  K   S  G  L  V
1101 TCGACAAGCG CCCGAACCCC AACGACGGCC GCGGCACGCT CGCCTCCATC ACGGAGAAGG GCCGCGAGGT CGTCGAGGCG GCCACCCGCG AGCTGATGGC
       D  K  R   P  N  P   N  D  G  R   G  T  L   A  S  I   T  E  K  G   R  E  V   V  E  A   A  T  R  E   L  M  A
1201 GATGGACTTC GGGCTCGGGG TGTACGACGC GGAGGAGTGC GGGGAGATCT TCGCGATGCT GCGGCCCCTG CGGGTGGCGG CGCGCGATTT CGAGGAGCAG
       M  D  F   G  L  G  V   Y  D  A   E  E  C   G  E  I  F   A  M  L   R  P  L   R  V  A  A   R  D  F   E  E  Q
1301 TAGGGCCCGC CCGGTGAGAA GTGGGATCGG GTCGTCCCGG TACGGGCGGG GGCGGCGAAG ATCGCGTGAA AAGGGCGGTT ACGCTCGTAG CCATGAAACG
     *                                                                                    orf 2 ->       M  K  R
1401 CAGCGTGCTG ACCCGCTACC GGGTGATGGC CTACGTCACC GCCGTCATGC TCCTCATCCT GTGCGCCTGC ATGGTGGCCA AGTACGGCTT CGACAAGGGC
       S  V  L   T  R  Y  R   V  M  A   Y  V  T   A  V  M  L   L  I  L   C  A  C   M  V  A  K   Y  G  F   D  K  G
1501 GAGGGTCTGA CCCTCGTCGT GTCGCAGGTG CACGGCGTGC TCTACATCAT CTACCTGATC TTCGCCTTCG ACCTGGGCTC CAAGGCGAAG TGGCCGTTCG
      E  G  L  T   L  V  V   S  Q  V   H  G  V  L   Y  I  I   Y  L  I   F  A  F  D   L  G  S   K  A  K   W  P  F  G
1601 GCAAGCTGCT CTGGGTGCTG GTCTCGGGCA CGATCCCGAC CGCCGCCTTC TTCGTCGAGC GCAAGGTCGC CCGTGACGTC GAGCCGCTGA TCGCCGACGG
       K  L  L   W  V  L   V  S  G  T   I  P  T   A  A  F   F  V  E  R   K  V  A   R  D  V   E  P  L  I   A  D  G
1701 CTCCCCGGTC ACCGCGAAGG CGTAACCCGC ACCGCCACGG ACAGGTCCGT GGCGGTTGGC CATCGACTTT TACTAGGACG TCCTAGTAAA TTCGATGGTA
       S  P  V   T  A  K  A   *                                                                  orf1 icmA ->  M
1801 TGGACGCTGA CGCGATCGAG GAAGGCCGCC GACGCTGGCA GGCCCGTTAC GACAAGGCCC GCAAGCGCGA CGCGGACTTC ACCACGCTCT CCGGGGACCC
       D  A  D   A  I  E   E  G  R  R   R  W  Q   A  R  Y   D  K  A  R   K  R  D   A  D  F   T  T  L  S   G  D  P
1901 CGTCGACCCC GTCTACGGCC CCCGGCCCGG GGACACGTAC GACGGGTTCG AGCGGATCGG CTGGCCGGGG GAGTACCCCT TCACCCGCGG GCTCTACGCC
       V  D  P   V  Y  G  P   R  P  G   D  T  Y   D  G  F  E   R  I  G   W  P  G   E  Y  P  F   T  R  G   L  Y  A
2001 ACCGGGTACC GCGGCCGCAC CTGGACCATC CGCCAGTTCG CCGGCTTCGG CAACGCCGAG CAGACGAACG AGCGCTACAA GATGATCCTG GCCAACGGCG
      T  G  Y  R   G  R  T   W  T  I   R  Q  F  A   G  F  G   N  A  E   Q  T  N  E   R  Y  K   M  I  L   A  N  G  G
2101 GCGGCGGCCT CTCCGTCGCC TTCGACATGC CGACCCTCAT GGGCCGCGAC TCCGACGACC CGCGCTCGCT CGGCGAGGTC GGCCACTGCG GTGTCGCCAT
       G  G  L   S  V  A   F  D  M  P   T  L  M   G  R  D   S  D  D  P   R  S  L   G  E  V   G  H  C  G   V  A  I
2201 CGACTCCGCC GCCGACATGG AGGTCCTCTT CAAGGACATC CCGCTCGGCG ACGTCACGAC GTCCATGACC ATCAGCGGGC CCGCCGTGCC CGTCTTCTGC
       D  S  A   A  D  M  E   V  L  F   K  D  I   P  L  G  D   V  T  T   S  M  T   I  S  G  P   A  V  P   V  F  C
2301 ATGTACCTCG TCGCGGCCGA GCGCCAGGGC GTCGACCCGG CCGTCCTCAA CGGCACGCTG CAGACCGACA TCTTCAAGGA GTACATCGCC CAGAAGGAGT
      M  Y  L  V   A  A  E   R  Q  G   V  D  P  A   V  L  N   G  T  L   Q  T  D  I   F  K  E   Y  I  A   Q  K  E  W
2401 GGCTCTTCCA GCCCGAGCCG CACCTGCGCC TCATCGGCGA CCTGATGGAG CACTGCGCGC GCGACATCCC CGCGTACAAG CCGCTCTCGG TCTCCGGCTA
       L  F  Q   P  E  P   H  L  R  L   I  G  D   L  M  E   H  C  A  R   D  I  P   A  Y  K   P  L  S  V   S  G  Y
2501 CCACATCCGC GAGGCCGGGG CGACGGCCGC GCAGGAGCTC GCGTACACCC TCGCGGACGG CTTCGGGTAC GTGGAACTGG GCCTCTCGCG CGGCCTGGAC
       H  I  R   E  A  G  A   T  A  A   Q  E  L   A  Y  T  L   A  D  G   F  G  Y   V  E  L  G   L  S  R   G  L  D
2601 GTGGACGTCT TCGCGCCCGG CCTCTCCTTC TTCTTCGACG CGCACGTCGA CTTCTTCGAG GAGATCGCGA AGTTCCGCGC CGCACGCCGC ATCTGGGCGC
      V  D  V  F   A  P  G   L  S  F   F  F  D  A   H  V  D   F  F  E   E  I  A  K   F  R  A   A  R  R   I  W  A  R
2701 GCTGGCTCCG GGACGAGTAC GGAGCGAAGA CCGAGAAGGC ACAGTGGCTG CGCTTCCACA CGCAGACCGC GGGGGTCTCG CTCACGGCCC AGCAGCCGTA
       W  L  R   D  E  Y   G  A  K  T   E  K  A   Q  W  L   R  F  H  T   Q  T  A   G  V  S   L  T  A  Q   Q  P  Y
2801 CAACAACGTG GTGCGGACGG CGGTGGAGGC CCTCGCCGCG GTGCTCGGCG GCACGAACTC CCTGCACACC AACGCTCTCG ACGAGACCCT TGCCCTCCCC
      N  N  V   V  R  T  A   V  E  A   L  A  A   V  L  G  G   T  N  S   L  H  T   N  A  L  D   E  T  L   A  L  P
2901 AGCGAGCAGG CCGCGGAGAT CGCGCTGCGC ACCCAGCAGG TGCTGATGGA GGAGACCGGC GTCGCCAACG TCGCGGACCC GCTGGGCGGC TCCTGGTACA
      S  E  Q  A   A  E  I   A  L  R   T  Q  Q  V   L  M  E   E  T  G   V  A  N  V   A  D  P   L  G  G   S  W  Y  I
3001 TCGAGCAGCT CACCGACCGC ATCGAGGCCG ACGCCGAGAA GATCTTCGAG CAGATCAGGG AGCGGGGGCG GCGGGCCTGC CCCGACGGGC AGCACCCGAT
       E  Q  L   T  D  R   I  E  A  D   A  E  K   I  F  E   Q  I  R  E   R  G  R   R  A  C   P  D  G  Q   H  P  I
3101 CGGGCCGATC ACCTCCGGCA TCCTGCGCGG CATCGAGGAC GGCTGGTTCA CCGGCGAGAT CGCCGAGTCC GCCTTCCAGT ACCAGCGGTC CCTGGAGAAG
       G  P  I   T  S  G  I   L  R  G   I  E  D   G  W  F  T   G  E  I   A  E  S   A  F  Q  Y   Q  R  S   L  E  K
3201 GGCGACAAGC GGGTCGTCGG CGTCAACTGC CTCGAAGGCT CCGTCACCGG CGACCTGGAG ATCCTGCGCG TCAGCCACGA GGTCGAGCGC GAGCAGGTGC
      G  D  K  R   V  V  G   V  N  C   L  E  G  S   V  T  G   D  L  E   I  L  R  V   S  H  E   V  E  R   E  Q  V  R
3301 GGGAGCTTGC GGGGCGCAAG GGGCGGCGTG ACGATGCGCG GGTGCGGGCC TCGCTCGACG CGATGCTCGC CGCTGCGCGG GACGGGTCGA ACATGATTGC
       E  L  A   G  R  K   G  R  R  D   D  A  R   V  R  A   S  L  D  A   M  L  A   A  A  R   D  G  S  N   M  I  A
3401 CCCCATGCTG GAGGCGGTGC GGGCCGAGGC GACCCTCGGG GAGATCTGCG GGGTGCTTCG CGATGAGTGG GGGGTCTACG TGGAGCCGCC CGGGTTCTGA
       P  M  L   E  A  V  R   A  E  A   T  L  G   E  I  C  G   V  L  R   D  E  W   G  V  Y  V   E  P  P   G  F  *
3501 GGGCGCGCTC CCTTTGCCTG CGGGTCTGCT GTGGCTGGTC GCGCAGTTCC CCGCACCCCT GAAAGACCCC GGCGCTTTCC CTTCCTGGCT CGCCTCGTCG
3601 CTGTCTGCGG GGCCGTGGGG GCTGGTCGCG CAGTTCCCCG CGCCCCTGCC CGCACCTGCG CCCCGCCGCC TGCATGCCGC CCCCACCCTG ACGGGGGCGT
3701 TCGGGGCCCA CCCTGACGGG TGCGGTCGGG GCGTGCCGGG GTCTTTTAGG GGCGCGGGGA ACTGCGCGAG CAACCCCCAC CCACCCGCAG GTGCACGCGG
3801 AGCGGCGGAC GCCCCGCAGA CGGGGGCAAA ACGGGCGGAG TGCCCCCGCC CGCCGGGCGG CGCGAATTCG TAGGTTTAAG GGGCAGGGGT CAGGGCAGGC
3901 GCCGAGCCGG TCAACCGCCC CCGTCCCAGG AGACCCCGTG ACCTCGACCG GCCACGCCCG CACCGCCGCC ATCGCCATCG GAGCCGCCAC CGCCACCGTC
                                orf 5 ->        V   T  S  T  G   H  A  R   T  A  A   I  A  I  G   A  A  T   A  T  V
4001 CTCGGCGCGC TGCTGGTCGG CGGCTCCGGC GAGGTGAGTG CGAGCCCGCC GCCCGAGCCC AAGGTCCAGG ACGACTTCGA CTCCCTCGGC CCCGAGGTGC
      L  G  A  L   L  V  G   G  S  G   E  V  S  A   S  P  P   P  E  P   K  V  Q  D   D  F  D   S  L  G   P  E  V  R
4101 GCGCCGCGAA GCTCTCCGAC GGGCGGACGG CCCACTACTC GGACACGGGC GACAAGGACG GCAAGCCGGC CCTGTTCATC GGCGGCACCG GCACGAGCGC
       A  A  K   L  S  D   G  R  T  A   H  Y  S   D  T  G   D  K  D  G   K  P  A   L  F  I   G  G  T  G   T  S  A
4201 CCGCGCCTCC CACATGACCG ACTTCTTCCG CTCGACGCGC GAGGACCTGG GCCTGCGCCT CATCTCCGTG GAGCGCAACG GCTTCGGCGA CACCGCGTTC
       R  A  S   H  M  T  D   F  F  R   S  T  R   E  D  L  G   L  R  L   I  S  V   E  R  N  G   F  G  D   T  A  F
4301 GACGAGAAGC TGGGCACCGC CGACTTCGCG AAGGACGCCC TCGAAGTCCT CGACCGGCTC GG
      D  E  K  L   G  T  A   D  F  A   K  D  A  L   E  V  L   D  R  L  -->
```

FIG. 4. **The nt sequence of the cloned DNA from *S. cinnamonensis*.** The location of orfs predicted from a FRAME analysis are shown. The *orf4* is encoded on the opposite strand (not shown), and the predicted aa sequence is shown *above* the nt sequence. The other orfs are encoded on the DNA strand shown, and the aa sequence is given below. The aa sequences of seven IcmA tryptic peptides, and the N-terminal sequence, determined by protein sequencing are given in *italics* and are *underlined*. The predicted start codons (determined for *orf1*) are shown in **bold**. An inverted repeat between *orf1* and *orf2* is *underlined*. The nt sequences at the 5′ and 3′ ends of *icmA* corresponding to the target sites of PCR primers KB3 and KB4 are also *underlined*.

was applied in portions to Ni^{2+}-NTA resin (3 × 1.5 cm, Pharmacia) and washed with buffer K, and His_6-IcmA was eluted with buffer L and then dialyzed against buffer M. Chromatography on MonoQ with buffer M and a gradient of 0–100% KCl (1 M) afforded His_6-IcmA (30 mg) which was homogeneous by SDS-PAGE (Fig. 7).

Plasmid pOCI633 was introduced into *S. lividans* 1362 (25). After growth in YEME (5 liters) with kanamycin (5 μg/ml) at 30 °C to an A_{600} of 0.7–1.0, the cultures were induced with Ts (5 μg/ml). After a further 12–15 h the cells were collected and sonicated in buffer K; cell debris was removed by centrifugation, and His_6-IcmA was purified as above (yield 16 mg).

Purification of IcmB from S. lividans

S. lividans 1362[pOCI633] was grown in YEME (5 liters) and induced with Ts, as described above. The cells were sonicated in buffer K; cell debris was removed; coenzyme B_{12} (10 μM) was added, and the protein (~1.6 g) was chromatographed in portions on Ni^{2+}-NTA resin (3 × 1.5 cm) in the dark. The resin was washed with buffer K containing coenzyme B_{12} (10 μM), before eluting with buffer L and dialyzing against buffer M (yield 24 mg). The sample was applied to MonoQ with buffer M containing coenzyme B_{12} (10 μM), washed with buffer M (no coenzyme B_{12}), and eluted with a gradient of 0–100% KCl (1 M) to give a fraction containing mainly His_6-IcmA (16 mg, eluting at 270 mM KCl), followed by a pink colored protein (2.5 mg, eluting at 370 mM KCl) with high ICM activity and a UV spectrum with an absorption maximum at 525 nm. The pink colored protein was dialyzed against buffer M, applied to MonoQ, and eluted in buffer M with a gradient of 0–100% KCl (1 M). The small subunit IcmB (~4 μg) eluted as a sharp peak at 200 mM KCl and was homogeneous by SDS-PAGE (see Fig. 7).

Purification of IcmB from S. cinnamonensis

S. cinnamonensis was grown for 3–4 days at 30 °C in YEME (5 liters) supplemented with valine (6.6 g/liter). The cells were collected and sonicated in buffer K, and cell debris was removed by centrifugation (yield 1.8 g of protein). To this was added recombinant His_6-IcmA (29 mg), prepared as described above, and coenzyme B_{12} (10 μM). The protein was chromatographed in portions on Ni^{2+}-NTA resin (3 × 1.5 cm) in the dark. The resin was washed with buffer K containing coenzyme B_{12} (10 μM), before eluting with buffer L and dialyzing against buffer M (yield 22 mg). The sample was chromatographed on MonoQ, as described above for IcmB from *S. lividans*. The protein fraction eluting at ~300 mM KCl (15 mg) contained mainly His_6-IcmA, whereas fractions containing the holoenzyme (4.5 mg) eluted at ~350 mM KCl and were dialyzed against buffer M. The protein was then applied to a gel filtration column (Superose 12, Pharmacia) and eluted at a flow rate of 0.2 ml/min with buffer N. A peak containing IcmB (~2 μg) eluted with an apparent mass of ~16–18 kDa (Fig. 7). This protein showed no ICM activity until both IcmA and coenzyme B_{12} were added (Fig. 8).

RESULTS

Enzyme Assay—The ICM assay involves hydrolyzing CoA thioesters at the end of the reaction, extraction of *n*- and isobutyric acids into ethyl acetate, and quantification by GC. Typical GC chromatograms from assays performed with recombinant ICM (see above) are shown in Fig. 8. To aid in the quantification of isobutyrate formed, a known amount of valeric acid was added to each assay as an internal standard.

Enzyme Purification and Peptide Sequencing—The ICM was purified as outlined in Table I. The MCM and ICM activities were separated by Q-Sepharose ion-exchange chromatography, with ICM eluting in the middle and MCM at the end of the salt gradient. The protein finally obtained was shown by SDS-PAGE to contain a major component with apparent mass ≈65 kDa, together with several minor components of lower mass (Fig. 2). Attempts to further purify the ≈65-kDa protein led to large losses in ICM activity. The ≈65-kDa protein was isolated by SDS-PAGE and subjected to N-terminal amino acid sequence analysis (see Fig. 4). The ≈65-kDa protein was also digested with trypsin, and the resulting peptides were analyzed by high pressure liquid chromatography and electrospray-mass spectrometry. The masses of tryptic fragments were compared with the MOWSE peptide mass fingerprint data base (33), but no similar entries were found. Several tryptic fragments were sequenced by Edman degradation (Fig. 4), revealing up to 75% sequence identity to segments of the MCM large subunit from *S. cinnamonensis* (6) and *P. shermanii* (7).

Table I
Purification of ICM from extracts of S. cinnamonensis

Purification step	Protein[a]	Isobutyrate[b]	n-Butyrate/isobutyrate[c]	Activity[d]
	mg	*μM*		*μmol/min/mg protein*
1. $[NH_4]_2SO_4$	ND			
2. DEAE	800	31	6.4:1	2.6×10^{-4}
3. Q-Sepharose	120	36	4.4:1	1.5×10^{-3}
4. Superdex	20	45	4.2:1	7.5×10^{-3}
5. Gel electrophoresis	2.5	53	4.3:1	8.9×10^{-3}
6. B_{12} affinity chromatography	0.26	80	2.6:1	2.3×10^{-2}

[a] Amount of protein available following each step, starting from ~500 g (wet weight) cell paste.
[b] The concentration of isobutyrate attained at the end of the assay is calculated by comparison to the amount of *n*-valeric acid in the EtOAc extract. In each assay, *n*-butyryl-CoA (250 μM) in buffer (200 μl) and coenzyme B_{12} were incubated with protein for 30 min at 30 °C (see "Experimental Procedures").
[c] The ratio of *n*- to isobutyrate peak volumes detected by GC is shown. This indicates how far the interconversion has proceeded in each assay. No correction for hydrolysis of substrates during the course of the assay is included.
[d] Activity indicates amount of isobutyrate formed (see Footnote *b*) in μmol/min/mg protein.

Gene Cloning and Sequencing—Two tryptic peptide sequences were used to design oligonucleotides for PCR. The PCR afforded a ≈310-bp DNA fragment, which was found to be 70% identical in DNA sequence and 55% in translated protein sequence to the MCM large subunit from *S. cinnamonensis*. This PCR product was used as a probe to isolate hybridizing clones from a genomic DNA library prepared in λEMBL4. From one clone, the region encoding the putative *icmA* gene was isolated and sequenced on both strands by the dideoxy method (see Figs. 3 and 4).

Sequence Analysis—The 4362-bp DNA segment sequenced showed a total G/C content of 71%. A frame analysis (Fig. 5) was performed using CODONPREFERENCE in the GCG software (The Genetics Computer Group, Madison, WI, version 8.1-UNIX (34)). This revealed three complete orfs (*orf1*, *orf2*, and *orf3* in Figs. 3 and 4), each with a G/C content of ≈75, ≈50, and ≈95% at the first, second, and third positions of each codon, respectively, which is highly characteristic of protein coding regions in *Streptomyces* DNA (35). Downstream of the presumptive stop codon of *orf1*, the G/C distribution changes (Fig. 5), strongly suggesting that the stop codon has been correctly identified. Two incomplete orfs (*orf4* and *orf5*) were also predicted, extending outward from each end of the region sequenced. The incomplete *orf4* (nt 1–583) shared over the available protein sequence a similarity of ≈32% to endoglucanases in the EMBL/SWISSPROT data base. Comparisons of *orf2* (nt 1393–1725), *-3* (nt 794–1303), and *-5* (nt 3938–4326) with the data base failed to identify proteins with significant sequence similarities, so their functions are presently uncertain. The *orf1* was identified as the putative *icmA* gene due to its high sequence similarity, at the DNA and protein levels, to the large

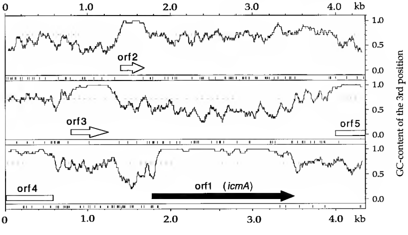

Fig. 5. **FRAME analysis performed using CODONPREFERENCE in the GCG software (34) (PrefWindow: 25, Rare codon Treshold: 0.1, BiasWindow: 25, density: 143.1).** The analysis shows the percentage G/C *versus* A/T at the third position of each codon for all possible reading frames. The location of orfs 1–5 deduced in this way (see text) is shown *below* each trace.

TABLE II
Protein sequence identities and similarities between IcmA and the large (MutB) and small (MutA) subunits of MCM from S. cinnamonensis and P. shermanii, and the human and mouse MCMs (using BESTFIT, in the GCG software)

Protein sequence	Identity	Similarity
	%	%
MCM *S. cinnamonensis* MutA	29.6	51.5
MCM *S. cinnamonensis* MutB	44.0	63.7
MCM *P. shermanii* MutA	26.1	51.1
MCM *P. shermanii* MutB	42.9	65.3
MCM mouse	40.4	63.7
MCM human	42.3	64.3

subunits of *S. cinnamonensis* and *P. shermanii* MCM, as well as to the human and mouse MCMs (Table II). The 3′-untranslated region downstream of *orf1* shows no homology at the nt or aa levels to MCMs, again consistent with the correct identification of the stop codon of *orf1*.

The N-terminal amino acid sequence determined for IcmA agrees with that predicted by the DNA sequence, starting at nt 1800 with an ATG codon. Termination occurs at nt 3500 with a TGA codon, corresponding to a protein with 566 aa, and a mass of 62,487 Da, which agrees well with the mass of ~65 kDa estimated by SDS-PAGE. The peptide sequences determined from tryptic fragments are encoded at the expected locations in the *icmA* gene sequence (Fig. 4). A comparison of the IcmA protein sequence, with those of the homodimeric and heterodimeric MCM large subunits from various organisms, was performed with PILEUP in the GCG software (Fig. 9). A DOT-PLOT comparison between ICM and the MCM large subunit from *P. shermanii* is shown in Fig. 10.

Disruption of the S. cinnamonensis icmA Gene—A targeted insertional inactivation of the *icmA* gene in *S. cinnamonensis* was achieved by first inserting a cassette containing a functional Hm resistance gene (*hygB*) into the unique *Sac*I site within the cloned *S. cinnamonensis icmA* gene (Figs. 3 and 6). The *icmA* containing *hygB* was cloned into the vector pGM160 (29) to give plasmid pOCI643 which, however, could not be introduced into *S. cinnamonensis*, possibly due to instability of the plasmid under the growth conditions. Subsequently, by removing the entire *E. coli* sequences from pOCI643, and introducing a plasmid denaturation step (30), *S. cinnamonensis* Ts^R Hm^R transformants were isolated, which after further growth at 39 °C yielded Ts^S Hm^R colonies. A Southern blot hybridization analysis of genomic DNA isolated from one of these clones confirmed that the *icmA* gene had been inactivated, consistent with a double crossover event (Fig. 6). Extracts of the *S. cinnamonensis icmA::hygB* mutant were devoid of ICM activity.

Expression of the icmA Gene—The *icmA* gene was amplified by PCR using oligonucleotide primers incorporating *Nde*I and *Bam*HI sites, such that the *Nde*I site incorporates an ATG start codon. The PCR product was cloned after digestion with *Nde*I and *Bam*HI between the *Nde*I/*Bam*HI sites in pET3a to afford pOCI614. After introduction into *E. coli* BL21(DE3)pLysS and induction with IPTG at 30 °C, large amounts of soluble protein were isolated, with the correct apparent mass on SDS-PAGE, and the correct N-terminal amino acid sequence. This protein, however, was devoid of ICM activity.

To produce IcmA in *S. lividans* 1326, the gene was cloned on the *Nde*I/*Bam*HI fragment into the high copy number expression vector pIJ4123 to afford pOCI633. Only a very low ICM activity was found in cell extracts of *S. lividans* 1326[pIJ4123] grown in YEME. However, cell extracts from *S. lividans* 1326[pOCI633] after induction with Ts showed high levels of ICM activity, typically about 5–10 × higher than seen in extracts of *S. cinnamonensis*. The His_6-IcmA was purified to homogeneity by Ni^{2+}-chelate affinity and gel filtration chromatography but showed no mutase activity (Fig. 8). The same His_6-IcmA was also produced in *E. coli* using the vector pET14b (Novagen) but again showed no ICM activity.

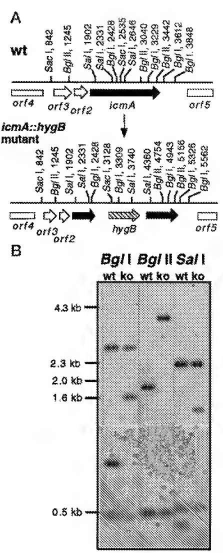

FIG. 6. *A*, the genomic region of wild type (*wt*) *S. cinnamonensis* encoding IcmA, as deduced from the sequenced DNA (see Figs. 3 and 4). The restriction sites (*Bgl*I, *Bgl*II, *Sal*I, and *Sac*I) and their positions found within the 4362 bp sequenced are indicated. Below is a map showing the expected organization of orfs, and positions of restriction sites, after a double crossover from the *hygB*-containing *icmA* gene in the vector and selection for Ts^S and Hm^R. *B*, hybridization analysis of a Southern blot of DNA from the Ts^S Hm^R transformant (denoted *ko*) and the wild type strain, digested with *Bgl*I, *Bgl*II, or *Sal*I. The probe was the PCR fragment amplified by primers KB3 and KB4 (see "Experimental Procedures") and digested with *Nde*I/*Bam*HI. This probe includes the entire *orf1*. The positions of size markers are shown to the *left*.

Purification of an ICM Small Subunit—Cell extracts from *S. lividans* 1326[pOCI633] were fractionated by metal-chelate affinity chromatography in the presence of coenzyme B_{12}, to recover His_6-IcmA and its associated subunit. Subsequent ion-exchange chromatography on MonoQ (Pharmacia) in the absence of coenzyme B_{12} gave a protein similar to 17 kDa by SDS-PAGE (Fig. 7), which by itself was devoid of ICM activity, but gave highly active ICM after incubation with His_6-IcmA (or IcmA) and coenzyme B_{12}. The intact holoenzyme showed a UV-visible absorption spectrum with a maximum at 525 nm typical of protein-bound adenosyl cobalamin (data not shown).

The Journal of Biological Chemistry

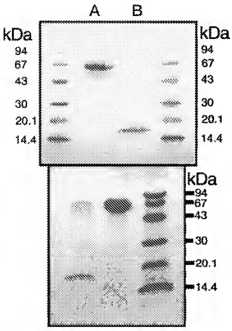

FIG. 7. **Above is shown a Coomassie Blue-stained SDS-PAGE (8–25% gradient) gel of purified His_6-IcmA from *E. coli* (*lane A*) and IcmB isolated from *S. lividans* (*lane B*).** Below is shown a Coomassie Blue-stained SDS-PAGE (20% homogeneous) gel with purified His_6-IcmA (*lane A*) and IcmB isolated from *S. cinnamonensis* (*lane B*). The positions of size markers are indicated.

The His_6-IcmA alone showed no UV-visible absorption maximum at 525 nm and does not bind coenzyme B_{12} under these conditions.

The ICM small subunit was isolated from wild type *S. cinnamonensis*, and from the *icmA::hygB* mutant (see above), in a similar way by addition of His_6-IcmA to cell extracts, followed by metal-chelate affinity chromatography in the presence of coenzyme B_{12}, ion-exchange chromatography on MonoQ, and gel filtration. The yield of the small subunit was lower, but SDS-PAGE again revealed a protein of ~17 kDa (Fig. 7), which by itself was inactive but yielded highly active ICM upon incubation with both coenzyme B_{12} and His_6-IcmA (Table III and Fig. 8), and afforded a holoenzyme with a UV-visible maximum at 525 nm.

DISCUSSION

Crucial to any enzyme purification is an assay that allows detection and quantification of catalytic activity. The assay for ICM used here is sensitive but ill-suited for accurate quantification of specific activity, especially when limited amounts of protein are available. For a typical assay during the purification of ICM, sufficient protein (~50–200 μg) was taken to afford between a ≈10:1 to 2:1 ratio (as determined by GC) of *n*- to isobutyrates in a single 30-min incubation at 30 °C, with *n*-butyryl-CoA as substrate. The amount of isobutyrate formed per min per mg of protein was then estimated, based on this single time point in the reaction. This gives an estimate of the mutase activity at each stage of the purification (Table I) but clearly does not correspond to the specific activity of the enzyme.

The enzyme is present in low amounts in cell extracts of *S. cinnamonensis* but is stable at room temperature over several hours. A variety of chromatographic methods failed to yield a significant improvement in purity without incurring major losses of ICM activity. With hindsight, it seems likely that these losses were due to the separation of subunits of the enzyme. This was not anticipated, since we had succeeded in purifying the heterodimeric MCM from *S. cinnamonensis* without major difficulties.[2] However, a significant gain in ICM purity was achieved by incorporating vitamin B_{12} affinity chromatography late in the purification scheme (Table I).

After six purification steps the ICM contained a major component with apparent mass ≈65 kDa on SDS-PAGE, along with several proteins of lower mass (Fig. 2). No protein of

TABLE III
Estimations of ICM activity reconstituted by mixing recombinant His_6-IcmA (produced in E. coli) with IcmB isolated from S. cinnamonensis (see Fig. 7) in the presence of coenzyme B_{12} (20 μM)

Without added coenzyme B_{12} the enzyme is completely inactive. The protein concentrations were determined by the Bradford assay, assuming a mass of 17 kDa for IcmB. The total amount of each subunit added to each assay in picomoles is shown. The substrate was *n*-butyryl-CoA (280 μM) in potassium phosphate buffer (50 mM, pH 7.4) with EDTA (5 mM) and glycerol (5% v/v), at 30 °C. The *i*/*n* ratio gives the ratio of *iso*- to *n*-butyrate determined by GC after 30 min incubation (compare Fig. 8), from which the amount of isobutyryl-CoA formed and hence the enzyme activity were determined (see "Discussion"). Compare with Table I.

Assay	IcmA	IcmB	i/n ratio	Activity
	pmol	*pmol*		*μmol/min/mg*
1	2.5	5	0.12	1.0
2	2.5	10	0.09	0.78
3	5	10	0.22	0.85
4	5	2.5	0.10	0.82
5	10	2.5	0.13	1.04
6	10	5	0.29	1.05

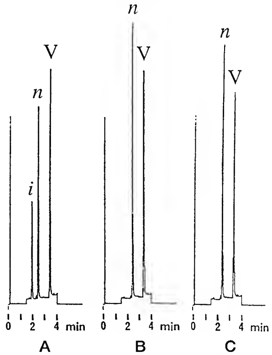

FIG. 8. **Gas chromatograms from assays of ICM activity performed with purified His_6-IcmA (~1 μg) and small subunit from *S. cinnamonensis* (~0.1 μg) with added coenzyme B_{12} (*A*); only the small subunit (~0.1 μg) with added coenzyme B_{12} (*B*); both proteins as in *A*, but without added coenzyme B_{12} (*C*).** Only in *A* is the formation of isobutyrate (*i*) from *n*-butyrate (*n*) apparent. The valeric acid added as a standard prior to extraction of the *n* and isobutyrate is denoted by *V*. The IcmA protein alone shows no activity and gives assay results exactly comparable to those shown in *B* and *C*.

[2] A. Leiser, unpublished work.

```
 human MLRAKNQLFLLSPHYLRQVKESSGSRLIQQRLLHQQQPLHPEWAALAKKQLKGKNPEDLIWHTPEGISIKPLYSKRD..TMDLPE..ELPGVKPFTRGPY   96
 mouse MLRAKNQLFLLSPHYLKQLNIPSASRW..KRLLHQQQPLHPEWAVLAKKQLKGKNPEDLIWHTPEGISIKPLYSRAD..TLDLPE..ELPGVKPFTRGPY   94
  pgin ............................MKPNYKDIDIKSAGFVAKDATRWAEEKGIVADWRTPEQIMVKPLYTKDDLEGMEHLD..YVSGLPPFLRGPY 70
psherm .......................MSTLPRFDSVDLGNAPVPADAARRFEELAAKAGTGEAWETAEQIPVGTLFNEDVYKDMDWLD..TYAGIPPFVHGPY   75
scinna ....................MRIPEFDDIELGAGGGPSGSAEQWRAAVKESVGKSESDLLWETPEGIAVKPLYTGADVEGLDFLE..TYPGVAPYLRGPY   78
   icm ..................................MDADAIEEGRRRWQARYDKARKRDADFTTLSGDPVDPVYGPRPGDTYDGFERIGWPGEYPFTRGLY   66
                                                                     *                           *  *   * *

 human PTMYTFRPWTIRQYAGFSTVEESNKFYKDNIKAGQQGLSVAFDLATHRGYDSDNPRVRGDVGMAGVAIDTVEDTKILFDGIPLEKMSVSMTMNGAVIPVL   196
 mouse PTMYTYRPWTIRQYAGFSTVEESNKFYKDNIKAGQQGLSVAFDLATHRGYDSDNPRVRGDVGMAGVAIDTVEDTKILFDGIPLEKMSVSMTMNGAVIPVL   194
  pgin SGMYPMRPWTIRQYAGFSTAEESNAFYRRNLASGQKGLSVAFDLATHRGYDADHSRVVGDVGKAGVSICSLEDMKVLFDGIPLSKMSVSMTMNGAVLPIL   170
psherm ATMYAFRPWTIRQYAGFSTAKESNAFYRRNLAAGQKGLSVAFDLPTHRGYDSDNPRVAGDVGMAGVAIDSIYDMRELFAGIPLDQMSVSMTMNGAVLPIL   175
scinna PTMYVNQPWTIRQYAGFSTAEESNAFYRRNLAAGQKGLSVAFDLPTHRGYDSDHPRVTGDVGMAGVAIDSIYDMRQLFDGIPLDKMTVSMTMNGAVLPVL   178
   icm ATGYRGRTWTIRQFAGFGNAEQTNERYKMILANGGGGLSVAFDMPTLMGRDSDDPRSLGEVGHCGVAIDSAADMEVLFKDIPLGDVTTSMTISGPAVPVF   166
          *    ***** ***      *  *      *  *******  *  * * *  *  * **  ** *    *   **  ***     ***  *   *

 human ANFIVTGEEQGVPKEKLTGTIQNDILKEFMVRNTYIFPPDPSMKIIADIFEYTAKHMPKFNSISISGYHMQEAGADAILELAYTLADGLEYSRTGLQAGL   296
 mouse ATFIVTGEEQGVPKEKLTGTIQNDILKEFMVRNTYIFPPEPSMKIIADIFQYTAQHMPKFNPISISGYHMQEAGTDAILELAYTIADGLEYCRTGLQAGL   294
  pgin AFYINAGLEQGAKLEEMAGTIQNDILKEFMVRNTYIYPPEFSMRIIADIFEYTSQNMPKFNSISISGYHMQEAGATADIEMAYTLADGMQYLKAGIDAGI   270
psherm ALYVVTAEEQGVKPEQLAGTIQNDILKEFMVRNTYIYPPQPSMRIISEIFAYTSANMPKWNSISISGYHMQEAGATADIEMAYTLADGVDYIRAGESVGL   275
scinna ALYIVAAEEQGVPPEKLAGTIQNDILKEFMVRNTYIYPPKPSMRIISDIFAYTSQKMPRYNSISISGYHIQEAGATADLELAYTLADGVEYLRAGQEAGL   278
   icm CMYLVAAERQGVDPAVLNGTLQTDIFKEYIAQKEWLFQPEPHLRLIGDLMEHCARDIPAYKPLSVSGYHIREAGATAAQELAYTLADGFGYVELGLSRGL   266
                **       ** * ** **          *      *           *     * ****  ***  *  * *** ***  *   *   *

 human TIDEFAPRLSFFWGIGMNFYMEIAKMRAGRRLWAHLIEKMFQPKNSKSLLLRAHCQTSGWSLTEQDPYNNIVRTAIEAMAAVFGGTQSLHTNSFDEALGL   396
 mouse TIDEFAPRLSFFWGIGMNFYMEIAKMRAGRRLWAHLIEKMFQPKNSKSLLLRAHCQTSGWSLTEQDPYNNIVRTAIEAMAAVFGGTQSLHSNSFDEALGL   394
  pgin DVDAFAPRLSFFWAIGVNHFMEIAKMRAARLLWAKIV.KSFGAKNPKSLALRTHSQTSGWSLTEQDPFNNVGRTCIEAMAAALGHTQSLHTNALDEAIAL   369
psherm NVDQFAPRLSFFWGIGMNFFMEVAKLRAARMLWAKLV.HQFGPKNPKSMSLRTHSQTSGWSLTAQDVYNNVVRTCIEAMAATQGHTQSLHTNSLDEAIAL   374
scinna DVDAFAPRLSFFWAIGMNFFMEVAKLRAARLLWAKLV.KQFDPKNAKSLSLRTHSQTSGWSLTAQDVFNNVTRTCVEAMAATQGHTQSLHTNALDEALAL   377
   icm DVDVFAPGLSFFFDAHVDFFEEIAKFRAARRIWARWLRDEYGAKTEKAQWLRFHTQTAGVSLTAQQPYNNVVRTAVEALAAVLGGTNSLHTNALDETLAL   366
         * *** ****         * ** ** *  **         *  *   ** * ** * *** *   **  **  ** **  * * *** *  **   *

 human PTVKSARIARNTQIIIQEESGIPKVADPWGGSYMMECLTNDVYDAALKLINEIEEMG................GMAKAVAEGIPKLRIEECAARRQARID   480
 mouse PTVKSARIARNTRIIIQEESGIPKVADPWGGSYMMESLTNDVYEAALKLIYEVEEMG................GMAKAVAEGIPKLRIEECAARRQARID   478
  pgin PTDFSARIARNTQIYIQEETLVCKEIDPWGGSYYVESLTNELVHKAWTLIKEVQEMG................GMAKAIETGLPKLRIEEAAARTQARID   453
psherm PTDFSARIARNTQLFLQQESGTTRVIDPWSGSAYVEELTWDLARKAWGHIQEVEKVG................GMAKAIEKGIPKMRIEEAAARTQARID   458
scinna PTDFSARIARNTQLLIQQESGTTRTIDPWGGSAYVEKLTYDLARRAWQHIEEVEAAG................GMAQAIDAGIPKLRVEEAAARTQARID   461
   icm PSEQAAEIALRTQQVLMEETGVANVADPLGGSWYIEQLTDRIEADAEKIFEQIRERGRRACPDGQHPIGPITSGILRGIEDGWFTGEIAESAFQYQRSLE   466
       *    * **  *      *       **  **   * **      *          *                *       *       * *   *

 human SGSEVIVGVNKYQLEKEDAVEVLAIDNTSVRNRQIEKLKKIKSSRDQALAEHCLAALTECAASG....DGNILALAVDASRARCTVGEITDALKKVFGE   575
 mouse SGSEVIVGVNKYHLEKEDSVHLLAIDIISLRKKQIEKLKKIKSSRDQALAEQCLSALTQCAASG....DGNILALAVDAARARCTVPEITDAFKKVFGE   573
  pgin SHQQVIVGVNKYRLPKEDPIDILEIDNTAVRKQQIERLNDLRSHRDEKAVQEALEAITKCVETK....EGNLLDLAVKAAGLRASLGEISDACEKVVGR   548
psherm SGRQPLIGVNKYRLEHEPPLDVLKVDNSTVLAEQKAKLVKLRAERDPEKVKAALDKITWAAGNPDDKDPDRNLLKLCIDAGRAMATVGEMSDALEKVFGR   558
scinna SGRQPVIGVNKYRVDTDEQIDVLKVDNSSVRAQQIEKLRRLREERDDAACQDALRALTAAAERGPGQGLEGNLLALAVDAARAKATVGEISDALESVYGR   561
   icm KGDKRVVGVNCLEGSVTGDLEILRVSH.EVEREQVRELAGRKGRRDDARVRASLDAMLAAARDG.....SNMIAPMLEAVRAEATLGEICGVLRDEWGV   559
              ***            *          *   *      **       *                 *       *        *         *

 human HKANDRMVSGAYRQEYGESKEITSAIKRVHKFMEREGRRPRLLVAKMGQDGHDRGAKVIATGFADLGFDVDIGPLFQTPREVAQQAVDADVHAVGVSTLA   675
 mouse HKANDRMVSGAYRQEFGESKEITSAIKRVNKFMEREGRRLGLLVAKMGKDGHDRGAKVIATGFADLGFDVDIGPLFQTPREVAHDAVDADVHAVGVSTHA   673
  pgin YKAVIRTISGVYSSESGEDKDFAHAKELAEKFAKKEGRQPRIMIAKMGQDGHDRGAKVVATGYADCGFDVDMGPLFQTPEEAARQAVENDVHVMGVSSLA   648
psherm YTAQIRTISGVYSKEVKNTPEVEEARELVEEFEQAEGRRPRILLAKMGQDGHDRGQKVIATAYADLGFDVDVGPLFQTPEETARQAVEADVHVVGVSSLA   658
scinna HAGQIRTISGVYRTEAGQSPSVERTRALVDAFDEAEGRRPRILVAKMGQDGHDRGQKVIASAFADLGFDVDVGPLFQTPAEVARQAVEADVHIVGVSSLA   661
   icm YVEPPGF*............................................................................................   566
                                                        ●●●●●●

 human AGHKTLVPELIKELNSLGRPDILVMCGGVIPPQDYEFLFEVGVSNVFGPGTRIPKAAVQVLDDIEKCLEKKQQSV*    750
 mouse AGHKTLVPELIKELTALGRPDILVMCGGVIPPQDYEFLYEVGVSNVFGPGTRIPRAAVQVLDDIEKCLAEKQQSV*    748
  pgin AGHKTLIPQVIAELEKLGRPDILVTAGGVIPAQDYDFLYQAGVAAIFGPGTPVAYSAAKVLEILLEE*......     715
psherm GGHLTLVPALRKELDKLGRPDILITVGGVIPEQDFDELRKDGAVEIYTPGTVIPESAISLVKKLRASLDA*....    728
scinna AGHLTLVPALREELAAEGRDDIMIVVGGVIPPQDVEALHEAGATAVFPPGTVIPDAAHDLVKRLAADLGHEL*..    733
   icm ...........................................................................
```

FIG. 9. **A PILEUP comparison of the aa sequences of (in descending order) the human and mouse MCMs, the large subunits of MCM from *Porphyromonos gingivalis*, *P. shermanii*, and *S. cinnamonensis*, and the ICM large subunit IcmA from *S. cinnamonensis*.** Residues that are conserved in all six sequences are indicated under the sequences by *. The D*X*H*XX*G motif is indicated by ●●● (see text).

higher mass was apparent by SDS-PAGE. Tryptic peptide fragments isolated from the ≈65-kDa protein showed high sequence identities (25–75%) to portions of the large subunits of MCM from both *S. cinnamonensis* (6) and *P. shermanii* (7), consistent with this being a subunit of a closely related enzyme.

A PCR-based reverse genetic approach then allowed the cloning and sequencing of the *icmA* gene (denoted *orf1* in Fig. 3). The translated sequence of 566 aa (M_r 62,487) shows a high similarity across almost its entire length to the large subunit of MCM from microbial sources (Fig. 10), as well as to the homodimeric human and mouse MCMs (Table II). It is noteworthy that an orf similar in size and sequence to that of the small subunit of MCMs from *S. cinnamonensis* and *P. shermanii* was not found directly adjacent to this *icmA* gene (Fig. 3). In contrast, the orfs for the large and small subunits of *S. cinnamonensis* and *P. shermanii* MCM possess overlapping stop and start codons, a device which is thought to lead to translational coupling and hence to the production of stoichiometric amounts of the two polypeptides.

Proof that *orf1* is necessary for ICM activity was obtained by disruption of the gene in *S. cinnamonensis*. By targeted insertional inactivation, the chromosomal *icmA* gene was replaced in a double crossover with a copy containing a functional Hm resistance gene inserted into its unique *Sac*I site (Fig. 6), using the vector pGM160. This vector contains a temperature-sensitive *Streptomyces* origin of replication, resulting in its loss from host cells grown at the non-permissive temperature of 39 °C (29). The vector has been used previously for gene disruptions in *S. cinnamonensis* (36, 37). The resulting *S. cinnamonensis icmA::hygB* mutant was devoid of ICM activity under the usual assay conditions, providing direct evidence for a functional role of *icmA* in ICM activity. This mutant will be of value to study the influence of ICM on polyketide antibiotic production in this strain, since the enzyme has been implicated in an important biochemical pathway, furnishing methylmalonyl-CoA from isobutyryl-CoA and *n*-butyryl-CoA (1).

A catalytic function for the IcmA protein was sought by expression in a heterologous host. In a first attempt, the protein was made in the cytoplasm of *E. coli*, by placing the *icmA* gene under the transcriptional control of a T7 RNA polymerase promoter in the plasmid pET3a (31). Although large amounts of soluble IcmA could be made in this way, it was devoid of ICM activity. The reason for the lack of mutase activity became clear after IcmA had been produced in *S. lividans*.

A second attempt to produce IcmA was made using a high copy number expression vector (pIJ4123) suitable for *Streptomyces* spp. (32). The vector contains the thiostrepton-inducible promoter and ribosome-binding site of the *tipA* gene. Immediately downstream is a translational start codon (ATG) followed by a sequence encoding a 20-residue N-terminal peptide including a His_6 tag and a thrombin recognition sequence, followed by a unique *Nde*I site allowing fusion of the peptide leader to the protein of interest. After subcloning the *icmA* gene into this vector, and introduction into *S. lividans* 1326, substantially

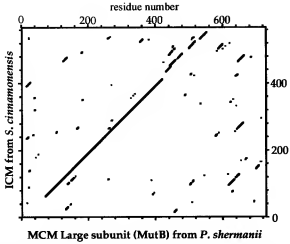

FIG. 10. **A DOTPLOT comparison of the MCM large subunit from *P. shermanii* and the IcmA from *S. cinnamonensis* (using COMPARE in the GCG software (34), window 30, stringency 15.0).**

higher ICM activity was detected in cell-free extracts than seen in *S. lividans* 1326 with pIJ4123. The His_6-IcmA was readily purified by Ni^{2+}-chelate affinity chromatography and gel filtration but then showed no ICM activity. Another protein fraction was detected, however, eluting from the gel filtration column after His_6-IcmA, which showed high ICM activity. This fraction contained several proteins in the size range 12–50 kDa, as well as small residual amounts of His_6-IcmA. As expected, the activity was dependent upon added coenzyme B_{12}. This suggested that at least one additional smaller subunit is necessary to complement the IcmA large subunit and afford active mutase *in vitro*. Indeed, it is notable that active mutase can be reconstituted with His_6-IcmA derived from *S. cinnamonensis* and small subunit(s) endogenous to the wild type *S. lividans* in which the large subunit had been produced.

The ICM small subunit was purified from *S. lividans* and subsequently also from *S. cinnamonensis*, by relying on its association with His_6-IcmA in the presence of coenzyme B_{12}, and exploiting the convenient His_6-affinity handle. The combination of metal-chelate affinity chromatography in the presence of coenzyme B_{12}, and subsequent ion-exchange and/or gel filtration chromatography in the absence of the coenzyme, gave a protein of about 17 kDa by SDS-PAGE (Fig. 7), which was inactive alone but afforded highly active ICM after addition of His_6-IcmA (or IcmA) and coenzyme B_{12} (Fig. 8). The activity of the reconstituted mutase was estimated to be approximately 1.0 μmol/min/mg, as shown in Table III. However, we note again here that the assay, as described above, is not well suited for determining specific activities. Nevertheless, this value can be compared with the activity determined for ICM isolated from *S. cinnamonensis*, as outlined in Table I. From this comparison it is clear that the mutase reconstituted from recombinant His_6-IcmA and IcmB from *S. cinnamonensis* (Table III and Fig. 7) has a higher activity than that found for the wild type enzyme at the end of the purification (Table I and Fig. 2).

A comparison of the IcmA protein sequence with those of the human and mouse MCMs and the large subunits from *P. gingivalis, S. cinnamonensis*, and *P. shermanii* MCMs performed using PILEUP in the GCG software (34) is shown in Fig. 9. A DOTPLOT comparison of IcmA and the MCM large subunit from *P. shermanii* is shown in Fig. 10. The DOTPLOT comparison reveals that the region of highest sequence similarity extends approximately over residues 60–400 in both proteins. The sequence identity in this region is about 50%. The most striking difference, however, is the significant truncation of IcmA in comparison to all MCMs (Fig. 9), corresponding to the loss of the C-terminal ≈160 amino acid residues from MCM. A second significant difference is a 16-residue insertion in IcmA (residues 424–439), which is absent in all the MCM sequences reported to date.

The crystal structure of the *P. shermanii* MCM reported recently (15) revealed an N-terminal $(\beta/\alpha)_8$-barrel domain in the large subunit, from residues A1–A400. The high sequence identity (≈50%) of this region to residues 1–392 in IcmA suggests that the $(\beta/\alpha)_8$-barrel is conserved in the structure of IcmA. Residues A401–A559 in the *P. shermanii* MCM correspond to a largely helical linker, which connects the $(\beta/\alpha)_8$-barrel with the C-terminal, so-called coenzyme B_{12} binding, flavodoxin-like domain (A560–A728). The linker residues A401–A559 in this MCM correspond in the sequence comparison to residues 393–560 in IcmA (Fig. 9), although the sequence identity is only ≈18% in this region (Fig. 10). But after just 6 more residues IcmA terminates.

A striking aspect of the recently determined crystal structures of MCM is the replacement of the dimethylbenzimidazole group of coenzyme B_{12} as an axial Co^{3+} ligand by the imidazole of a histidine situated in the C-terminal coenzyme B_{12} binding domain of the large subunit. This imidazole is linked through a hydrogen-bonded network to the side chains of two other residues forming a ligand triad (38), which in MCM is His^{610}-Asp^{608}-Lys^{604}. The nucleotide tail of the cofactor fills a cavity in this domain, which places the dimethylbenzimidazole group in a tight hydrophobic pocket. These intimate interactions between coenzyme B_{12} and protein suggest a key role for this domain in modulating the reactivity of MCM.

In the case of ICM, the large subunit contains no contiguous coenzyme B_{12} binding domain but requires a separate small subunit (IcmB) of ~17 kDa to bind coenzyme B_{12} and afford active mutase. This indicates that the IcmB small subunit has assumed the role of a coenzyme B_{12} binding domain in ICM and will most likely be homologous to the corresponding region of the MCM large subunit. In support of this conclusion, preliminary results from ongoing work have shown that the IcmB from *S. lividans* and *S. cinnamonensis* have N-terminal protein sequences that are about 70% identical to the coenzyme B_{12} binding domain in MCM (data not shown). In addition, a thorough sequence comparison has shown that the IcmB N-terminal protein sequence is not encoded in the genomic DNA shown in Fig. 4. Presently, we must conclude that the *icmB* gene is not encoded by one of the small orfs found adjacent to *icmA* in this work. Future work will focus on cloning *icmB* from *S. cinnamonensis*, the quaternary structure of the holoenzyme, and the determination of the kinetic and thermodynamic parameters of this mutase reaction.

Acknowledgments—We thank Dr. Mervyn Bibb for providing pIJ4123, Prof. Wohlleben for pGM160, and Annelies Meier for expert technical assistance.

REFERENCES

1. Reynolds, K. A., O'Hagan, D., Gani, D., and Robinson, J. A. (1988) *J. Chem. Soc. Perkin Trans. I* 3195–3207
2. Brendelberger, G., Retey, J., Ashworth, D. M., Reynolds, K., Willenbrock, F., and Robinson, J. A. (1988) *Angew. Chem. Int. Ed. Engl.* **27**, 1089–1090
3. Rétey, J. (1982) in *B12* (Dolphin, D., ed) Vol. 2, pp. 357–379, John Wiley & Sons, Inc., New York
4. Rétey, J., and Robinson, J. A. (1982) *Stereospecificity in Organic Chemistry and Enzymology*, Vol. 13, pp. 185–207, Verlag Chemie, Weinheim, Germany
5. Moore, B. S., Eisenberg, R., Weber, C., Bridges, A., Nanz, D., and Robinson, J. A. (1995) *J. Am. Chem. Soc.* **117**, 11285–11291
6. Birch, A., Leiser, A., and Robinson, J. A. (1993) *J. Bacteriol.* **175**, 3511–3519
7. Marsh, E. N., McKie, N., Davis, N. K., and Leadlay, P. F. (1989) *Biochem. J.* **260**, 345–352
8. Ledley, F. D., Lumetta, M., Nguyen, P. N., Kolhouse, J. F., and Allen, R. A.

(1988) *Proc. Natl. Acad. Sci. U. S. A.* **85,** 3518–3521
9. Jansen, R., Kalousek, F., Fenton, W., Rosenberg, L. E., and Ledley, F. D. (1989) *Genomics* **4,** 198–205
10. Wilkemeyer, M. F., Crane, A. M., and Ledley, F. D. (1990) *Biochem. J.* **271,** 449–455
11. Drennan, C. L., Huang, S., Drummond, J. T., Matthews, R. G., and Ludwig, M. L. (1994) *Science* **266,** 1669–1674
12. Stupperich, E., Eisinger, H. J., and Albracht, S. P. J. (1990) *Eur. J. Biochem.* **193,** 105–109
13. Marsh, E. N. G., and Holloway, D. E. (1992) *FEBS Lett.* **310,** 167–170
14. Padmakumar, R., Taoka, S., Padmakumar, R., and Banerjee, R. (1995) *J. Am. Chem. Soc.* **117,** 7033–7034
15. Mancia, F., Keep, N. H., Nakagawa, A., Leadlay, P. F., McSweeney, S., Rasmussen, B., Bösecke, P., Diat, O., and Evans, P. R. (1996) *Structure* **4,** 339–350
16. Bradford, M. M. (1976) *Anal. Biochem.* **72,** 248–254
17. Stark, W. M., Knox, N. G., and Westhead, J. E. (1968) *Antimicrob. Agents Chemother.* **1967,** 353–358
18. Priestley, N. D., and Robinson, J. A. (1989) *Biochem. J.* **261,** 853–861
19. Allen, R. H., and Majerus, P. W. (1972) *J. Biol. Chem.* **247,** 7695–7701
20. Yamada, R.-H., and Hogenkamp, H. P. C. (1972) *J. Biol. Chem.* **247,** 6266–6270
21. Yanisch-Perron, C., Vieira, J., and Messing, J. (1985) *Gene (Amst.)* **33,** 103–109
22. Agarwal, R. K., and Perl, A. (1993) *Nucleic Acids Res.* **21,** 5283–5284
23. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K. (1994) *Current Protocols in Molecular Biology*, John Wiley & Sons, Inc., New York
24. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
25. Hopwood, D. A., Bibb, M. J., Chater, K. F., Kieser, T., Bruton, C. J., Kieser, H. M., Lydiate, D. J., Smith, C. P., Ward, J. M., and Schrempf, H. (1985) *Genetic Manipulation of Streptomyces: A Laboratory Manual*, The John Innes Foundation, Norwich
26. Frischauf, A. M., Lehrach, H., Poutska, A., and Murray, N. (1982) *J. Mol. Biol.* **170,** 827–842
27. Sanger, F., Nicklen, S., and Coulson, A. (1977) *Proc. Natl. Acad. Sci. U. S. A.* **74,** 5463–5467
28. Lydiate, D. J., Ashby, A. M., Henderson, D. J., Kieser, H. M., and Hopwood, D. A. (1989) *J. Gen. Microbiol.* **135,** 941–955
29. Muth, G., Nussbaumer, B., Wohlleben, W., and Pühler, A. (1989) *Mol. Gen. Genet.* **219,** 341–348
30. Oh, S. H., and Chater, K. F. (1997) *J. Bacteriol.* **179,** 122–127
31. Studier, F. W., Rosenberg, A. H., Dunn, J. J., and Dubendorff, J. W. (1990) *Methods Enzymol.* **185,** 60–89
32. Takano, E., White, J., Thompson, C. J., and Bibb, M. J. (1995) *Gene (Amst.)* **166,** 133–137
33. Pappin, D. J. C., Hojrup, P., and Bleasby, A. J. (1993) *Curr. Biol.* **3,** 327–332
34. Devereux, J., Haeberli, P., and Smithies, O. (1984) *Nucleic Acids Res.* **12,** 387–395
35. Bibb, M. J., Findlay, P. R., and Johnson, M. W. (1984) *Gene (Amst.)* **30,** 157–166
36. Arrowsmith, T. J., Malpartida, F., Sherman, D. H., Birch, A., Hopwood, D. A., and Robinson, J. A. (1992) *Mol. Gen. Genet.* **234,** 254–264
37. Leiser, A., Birch, A., and Robinson, J. A. (1996) *Gene (Amst.)* **177,** 217–222
38. Ludwig, M. L., Drennan, C. L., and Matthews, R. G. (1996) *Structure* **4,** 505–512